MULTIPLE SCLEROSIS JOURNAL | MSJ

Systematic Review

Multiple Sclerosis Journal
2019, Vol. 25(14) 1848–1869
DOI: 10.1177/1352458519879094

Computerized neuropsychological assessment devices in multiple sclerosis: A systematic review

Curtis M Wojcik, Meghan Beier, Kathleen Costello, John DeLuca, Anthony Feinstein, Yael Goverover, Mark Gudesblatt, Michael Jaworski III, Rosalind Kalb, Lori Kostich, Nicholas G LaRocca, Jonathan D Rodgers and Ralph HB Benedict; National MS Society Cognition Work Team

Abstract
Background: The proliferation of computerized neuropsychological assessment devices (CNADs) for screening and monitoring cognitive impairment is increasing exponentially. Previous reviews of computerized tests for multiple sclerosis (MS) were primarily qualitative and did not rigorously compare CNADs on psychometric properties.
Objective: We aimed to systematically review the literature on the use of CNADs in MS and identify test batteries and single tests with good evidence for reliability and validity.
Method: A search of four major online databases was conducted for publications related to computerized testing and MS. Test–retest reliability and validity coefficients and effect sizes were recorded for each CNAD test, along with administration characteristics.
Results: We identified 11 batteries and 33 individual tests from 120 peer-reviewed articles meeting the inclusion criteria. CNADs with the strongest psychometric support include the CogState Brief Battery, Cognitive Drug Research Battery, NeuroTrax, CNS-Vital Signs, and computer-based administrations of the Symbol Digit Modalities Test.
Conclusion: We identified several CNADs that are valid to screen for MS-related cognitive impairment, or to supplement full, conventional neuropsychological assessment. The necessity of testing with a technician, and in a controlled clinic/laboratory environment, remains uncertain.

Keywords: Multiple sclerosis, computerized tests, cognition, systematic review, reliability, validity

Date received: 7 May 2019; revised: 4 September 2019; accepted: 6 September 2019.

Introduction

Roughly 50%–60% of patients with multiple sclerosis (MS) suffer from cognitive impairment (CI).[1–3] Currently, widespread screening for MS-related CI is limited by the considerable time and staff training required to administer traditional neuropsychological (NP) tests. Computerized neuropsychological assessment devices (CNADs) require fewer staff resources and less time to administer and score and may represent a viable alternative. This perspective hinges on both CNADs and conventional NP tests meeting adequate psychometric standards prior to clinical implementation.[4] Furthermore, confounding factors that may influence test validity should be accounted for, such as the degree of test automation and technician involvement/supervision. Previous reviews concluded that, despite some promising results, the scope of impairment targeted by CNADs is limited, motor confounders are not accounted for, and psychometric data are lacking.[5,6] While informative, the reviews did not cover psychometric properties in depth and omitted several individual tests not part of a larger battery.

Our objective was to assess the status of psychometric research on CNADs in MS and identify the most promising approaches for future research and clinical application. We endeavored to examine systematically the literature focusing specifically on test–retest reliability (consistency over repeated trials), ecological or

Correspondence to:
RHB Benedict
Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, State University of New York, 1010 Main Street, Buffalo, NY 14202, USA.
benedict@buffalo.edu

Curtis M Wojcik
Michael Jaworski III
Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, State University of New York, Buffalo, NY, USA

Meghan Beier
Division of Rehabilitation Psychology and Neuropsychology, Department of Physical Medicine and Rehabilitation, Johns Hopkins University School of Medicine, Baltimore, MD, USA

Kathleen Costello
Rosalind Kalb
Nicholas G LaRocca National Multiple Sclerosis Society, New York, NY, USA

John DeLuca
Department of Physical Medicine and Rehabilitation and Department of Neurology, Rutgers New Jersey Medical School, Newark, NJ, USA

Anthony Feinstein
Department of Psychiatry, University of Toronto, Toronto, ON, Canada

Yael Goverover
New York University, New York, NY, USA/South Shore Neurologic Associates, New York, NY, USA

Mark Gudesblatt
South Shore Neurologic Associates Islip, NY

Lori Kostich
The Mandell MS Center, Mount Sinai Rehabilitation Hospital, Hartford, CT, USA

Jonathan D Rodgers
Jacobs School of Medicine and Biomedical Sciences, University

predictive validity (relationships with real-world outcomes and general clinical measures), discriminant/known groups validity, and concurrent validity (correlations with established NP tests). Our overall aim was to help MS researchers and clinicians make informed decisions to apply select measures in MS research and clinical care.

Methods

Search protocol and inclusion

As per the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines,[7] we utilized PubMed, MEDLINE, PsycINFO, and Embase databases. The search terms and Boolean operators applied for each database were as follows: multiple sclerosis AND (neuropsychological test OR cognition OR cognitive function OR cognitive impairment OR memory OR processing speed OR executive function OR language OR attention OR visuospatial) AND (computer OR internet OR iPad OR tablet OR CNAD). Test names known to the authors were also employed: Automated Neuropsychological Assessment Metrics (ANAM), Central Nervous System-Vital Signs (CNSVS), Cognitive Drug Research (CDR) Battery, NeuroTrax, Cognitive Stability Index (CSI), Neurobehavioral Evaluation System (NES), Amsterdam Neuropsychological Test (ANT), Cambridge Neuropsychological Test Automated Battery (CANTAB), CogState Brief Battery (CBB), Cognivue, Cognistat, NeuroCog BAC App, CogniFit, BrainCheck, Standardized Touchscreen Assessment of Cognition (STAC), and the National Institutes of Health (NIH) Toolbox.

Inclusion criteria for the selected articles were as follows: (a) included an MS and healthy control (HC) sample, (b) included at least one measure with computerized stimuli and computerized response collection, (c) published in a peer-reviewed journal after 1990, and (d) included sufficient data to evaluate at minimum one of the following:

1. *Test/retest reliability*: Consistency of scores across multiple administrations using (a) Pearson (*r*) or intraclass correlation (ICC), (b) index of stability comparing scores between two or more administrations, or (c) if otherwise qualifying test development literature shows test–retest reliability in an HC sample.
2. *Discriminant or known groups validity*: MS and HC compared by (a) statistical test for difference between means, (b) means and standard deviations provided for both groups, (c) sensitivity/specificity values, or (d) area under the curve (AUC).
3. *Ecological validity*: Association between test and relevant patient outcomes via (a) correlation with or prediction of employment status or (b) correlation with Expanded Disability Status Scale (EDSS)[8] or Multiple Sclerosis Functional Composite (MSFC).[9]
4. *Concurrent validity*: Comparing CNAD and conventional NP tests via (a) correlations, (b) partial correlations accounting for demographic variables, (c) impairment detection rates, or (d) regression analyses with conventional test(s) predicting CNADs or vice versa.

The only exclusion criterion was a failure to meet one or more of the inclusion criteria. To obtain additional information regarding test characteristics (e.g. administration mode, hardware, etc.), information from CNAD websites was reviewed or provided by the test owner upon inquiry by the authors.

Effect size calculation

We calculated effect sizes (Cohen's *d*) comparing MS and HCs using available data or by converting another reported metric (e.g. Hedge's *g*). The average discriminant effect size for CNADs was calculated for each cognitive domain.

Domain identification

Battery subtests and individual tests were classified as measuring specific cognitive domains based largely on author descriptions, and tests used in multiple studies with varying domain attribution (e.g. *N*-Back—both processing speed and working memory) were classified as measuring multiple domains. Simple reaction time tests were classified as measuring cognitive processing speed.

Results

The PRISMA search found 120 articles (Figure 1), covering 11 batteries (i.e. multiple distinct cognitive functions) and 33 individual tests. CNADs with multiple indices within a single domain were classified as individual tests. For example, the Test of Attentional Performance (TAP) has multiple subtests, but each measures attention, so it was not classified as a "battery." Certain subtests were not included in the review, as they were never used in any study with an MS sample (e.g. several subtests of the CANTAB, ANAM, and TAP were not included as they were not used in a published MS study).

CNAD administration characteristics

Tables 1 and 2 reveal that most CNADs are Personal Computer (PC) based but several are dually compatible

at Buffalo, State University of New York, Buffalo, NY, USA/ Jacobs Neurological Institute, Buffalo, NY, USA/Canisius College, Buffalo, NY, USA

Ralph HB Benedict
Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, State University of New York, Buffalo, NY, USA/Jacobs Neurological Institute, Buffalo, NY, USA

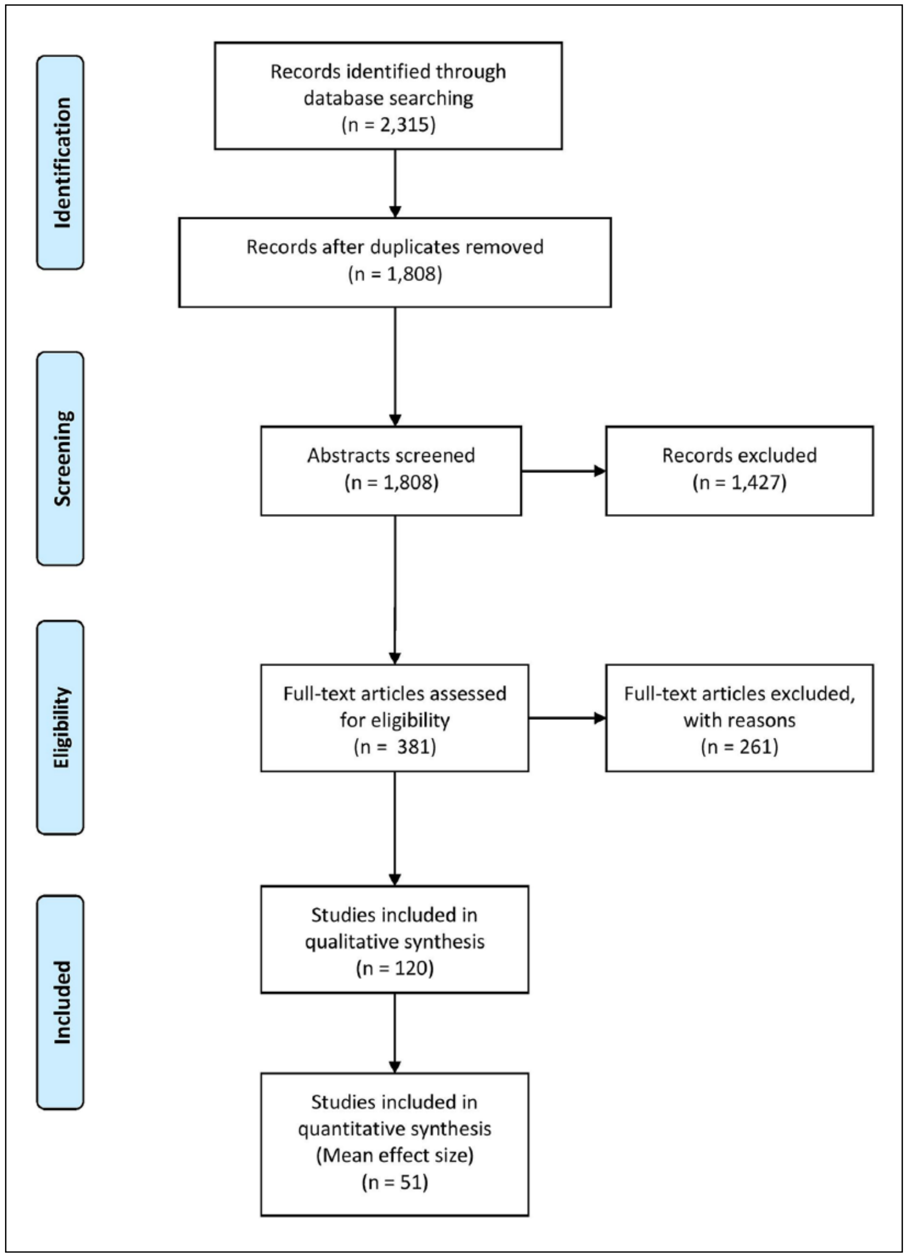

Figure 1. PRISMA systematic review flow diagram.

with either PC or iPad®/tablet administration, as claimed by the authors, but only the CBB empirically established the equivalence of PC and iPad administrations.[55] Most batteries were presented as requiring a technician for administration, as in NeuroTrax, CSI, NES2, CANTAB, ANAM, ANT, and CNSVS. The influence of technician supervision was investigated in only a few studies, results supporting the idea that human-supervised administration yields equivalent data to self-administration,[40,56,57] although this conclusion is deemed preliminary and likely to vary by test. Of the 11 CNAD batteries, 9 included a large normative database from which scores could be standardized. Few individual tests not part of a battery had large normative databases for standardized scores, but most were administered in at least one study that included HCs.

Psychometric findings by cognitive domain

Table 3 lists CNADs by functional domain (standalone or part of a larger battery) and shows results for the four psychometric properties. Supplementary Tables A and B provide more details regarding samples, effect sizes, and validity coefficients from each study.

Table 1. Test administration characteristics.

Battery	Administration		Time to administer	Alternate forms	Number of MS-related publications	Normative sample
	Mode	Technician				
Cognitive Drug Research (CDR) Battery[10]	PC	Self or Tech	15–20 minutes	Yes, at least 30 for dementia battery[11]	1	Several
NeuroTrax[12]	PC	Tech	45 minutes for full battery or shorter depending subtests selected	Yes, 3 equivalent	23	N = 1569, age range: 8–93, 58% female
Cognitive Stability Index (CSI)[13]	PC	Tech	25–35 minutes	Yes, 5 to 6 depending on subtest	1	N = 284, age range: 18–87, 53% female
Cambridge Neuropsychological Test Automated Battery (CANTAB)[14]	iPad	Tech	Depends on tests selected	Yes, but equivalence not yet established	8	Several
Neurobehavioral Evaluation System 2 (NES2)[15]	PC	Tech	20+ minutes	Unknown	1	N = 67, mean age = 56.5, 57.6% female[a]
Automated Neuropsychological Assessment Metrics (ANAM)[16]	PC	In MS, performance is better on some subtests in at-home, webcam-guided session vs live technician guidance	20–25 minutes	Yes, multiple alternate forms for each test	6	N = 107,671, age range: 17–65, 90% male
Amsterdam Neuropsychological Test (ANT)[17]	PC	Tech	Depends on tests selected	Yes, 3	3	Various normative samples depending on the subtest; ranges N = 14 to N = 330
Cognivue[18]	PC	Self	10 minutes	No	1	N = 401, age range: 55–95
CogniFit (CAB Battery)[19]	iPad/PC	Self	30–40 minutes	Unknown	1	N = 861, mean age = 65.7
CogState Brief Battery (CBB)[20]	iPad/PC	Self	20 minutes	Yes, stimuli can be randomized	5	N > 2500, age range: 18–99
Central Nervous System-Vital Signs (CNSVS)[21]	iPad/PC	Tech	Depends on tests selected	Except for memory tests, randomization allows for unlimited alternate forms	3	N > 1600, age range: 8–90

Reported sample details are those currently available for each test.
[a]Healthy sample from a research study, not currently used for standardization of scores.

Table 2. Test administration characteristics.

Test	Administration		Time to Administer	Alternate Forms	Number of MS-related Publications	Normative Sample
	Mode	Technician				
Computerized Test of Information Processing (CTIP)[22]	PC	Self	15 minutes	No, but no practice effects observed	8	N = 386, age range: 15–74
Computerized Stroop Test[23]	PC	Self	Varies, depending on modification of the task	Stimuli can be customized	8	Several[a]
Computerized Symbol Digit Modalities Test (C-SDMT)[24]	PC	Tech	Varies, ~90 seconds	Yes, randomizes key pairings	7	Several[a]
fMRI-SDMT[25]	fMRI	Tech	5 minutes	Unknown	2	1. N = 18, mean age = 41.14, 44% female[a] 2. N = 16, mean age = 19.1, 69% female[a]
Auto-SDMT[26]	PC	Self	Varies depending on participant	Fixed and variable key pairings	1	N = 32, mean age = 43.13, 72% female[a]
Computerized Incidental Learning Test (c-ILT)[27]	PC	Tech	Unknown	Yes, 2	1	N = 38, mean age = 54.3, 66% female[a]
Computerized Speed Cognitive Test (CSCT)[28]	PC	Tech	90 seconds	Yes, randomizes key pairings	4	N = 415, age range: 18 to >65
Paced Visual Serial Addition Test (PVSAT)[29]	PC	Tech	2–3 minutes	Yes, 2	5	Several[a]
Visual Threshold Serial Addition Task (VT-SAT)[30]	PC	Tech	3 minutes (depending on subject response to first trial)	Unknown	4	N = 32, mean age = 44[a]
N-Back (0-Back, 1-Back, and 2-Back)[31]	PC	Self	Varies depending on modification of the task	Stimuli can be customized	17	Several[a]
Janculjak Tests[32]	PC	Tech	Unknown	No	1	N = 42, mean age = 37, 55% female[a]
Computerized Paired Associate Learning (CPAL) Test[23]	PC	Tech	Unknown	Unknown	1	N = 40, mean age = 44.9, 62.5% female[a]
Sternberg Memory Scanning Task[33]	PC	Self	Varies, can be modified	Yes, can customize stimuli	2	Several
Computerized Salthouse Operational Working Memory Test/Keeping Track Task (KTT)[34]	PC	Tech	Varies depending on participant response speed	Yes, can customize	4	1. N = 35, mean age = 38, 74% female[a] 2. N = 36, mean age = 43.62[a]
Attention Network Test[35]	PC	Self	30 minutes	Yes, can customize stimuli	7	Several[a]
fMRI Tower of London Test[36]	fMRI	Tech	Unknown	Unknown	1	N = 18, mean age = 36.6, 33% female[a]
Denney Tests[37]	PC	Unknown	60–75 minutes	Unknown	1	N = 40, mean age = 40.3, 90% female[a]
Kujala Tests[38]	PC	Tech	60–120 minutes	No	1	N = 35, mean age = 43.5, 51% female[a]

(Continued)

Table 2. (Continued)

Test	Administration		Time to Administer	Alternate Forms	Number of MS-related Publications	Normative Sample
	Mode	Technician				
Graded Conditional Discrimination Tasks (GCDT)[39]	Visual Display Unit monitor with a sloping front panel with two push button switches	Unknown	30 minutes	Unknown	1	$N=25$, mean age = 44.1, 56% female[a]
The Processing Speed Test (PST)[40]	iPad	Self or Tech	2 minutes	Yes, randomizes key pairings	3	1. $N=116$, mean age = 42.5, 53% female[a] 2. $N=85$, mean age = 43.1, 75% female[a] 3. $N=24$, mean age = 44, 88% female[a]
SRT & CRT[41]	PC	Self, with Tech present	Unknown, part of laboratory session lasting 10 minutes	No	2	$N=98$, mean age = 46.4, 65.3% female[a]
Foong Tests[42]	PC	Unknown	Unknown	Unknown	1	$N=20$, mean age = 40.55, 55% female[a]
Conners' CPT-II[43,44]	PC	Self with Tech present	7–8 minutes	Yes	2	CPT-II: 1920 CPT-III: 1400
Computerized Digit Span[45]	PC	Tech	Unknown	Unknown	1	$N=20$, mean age = 38.7[a]
Amieva Computerized Tests[46]	Touchscreen PC	Tech	Unknown	Unknown	2	$N=43$[a] $N=44$, mean age = 37.8, 68% female[a]
Test of Attentional Performance (TAP)[47]	PC	Self with Tech present	Unknown	Some subtests have randomized stimuli	5	Varies depending on subtest
Computerized Reading Span Task/ Automated Ospan[48]	PC	Self	Unknown, part of laboratory session lasting 1 hour	Yes, can customize stimuli	4	$N=5537$, age range: 17–35
NIH Toolbox (Cognitive Battery)[49]	iPad	Tech	Depends on tests selected	Only for the Picture Sequence Memory Test	1	$N=1038$, mean age = 49.1
Touch Panel Tests[50]	Large touch panel screen desktop	Unknown	Unknown	No	1	Two samples: 1. $N=40$, mean age = 42.7, 68% female[a] 2. $N=15$, mean age = 54.6, 80% female[a]
Urban DailyCog©[51]	Virtual reality environment	Self with Tech present	20 minutes	No	1	$N=22$, mean age = 37.8 years, 72.7% female[a]
Actual Reality Task[52]	Virtual reality environment	Self with Tech present	10–20 minutes	No	3	$N=18$, mean age = 44.1, 22.2% female[a] $N=32$, mean age = 43.8, 72% female[a]
The Memory and Attention Test (MAT)[53]	PC	Self with Tech present	Unknown	Unknown	2	$N=84$, mean age = 39.9, 65.5% female[a]
Visual Search Task[54]	PC (touchscreen)	Self with Tech present	15–20 minutes	Performed on E-Prime, can program to change stimuli	1	$N=40$, mean age = 36.3, 70% female[a]

fMRI: functional magnetic resonance imaging; SDMT: Symbol Digit Modalities Test; SRT: Simple Reaction Time; CRT: Cued Recall Test; CPT II: Continuous Performance Test II; CPT III: Continuous Performance Test III; NIH: National Institutes of Health.
Reported sample details are those currently available for each test.
[a]Healthy sample from a research study, not currently used for standardization of scores.

Table 3. Psychometric characteristics of computerized tests.

Functional domain	Battery/test name	Subtest	Comparator(s)	Test–retest reliability	Validity		
					Discriminant/ known groups	Ecological	Concurrent
Processing speed	C-SDMT	C-SDMT	SDMT	✓	✓	✓	✓[a]
	CSCT	CSCT	SDMT	✓	✓	✓	✓[a]
	NeuroTrax	IPS	SDMT, DSST, PASAT	✓	✓	✓	✓[a]
	CDR	Speed of Memory	SDMT, DSST, PASAT	✓	✓	✓	✓[a]
	PST	PST	SDMT	✓	✓		✓[a]
	CNSVS	Processing Speed	SDMT, DSST, PASAT	Mixed	✓	✓	✓[a]
	CTIP	SRT	SDMT, DSST, PASAT	Mixed	✓	✓	✓[a]
	CTIP	CRT	SDMT, DSST, PASAT	Mixed	Mixed		✓[a]
	CTIP	SSRT	SDMT, DSST, PASAT	Mixed	✓		✓[a]
	CBB	Detection	SDMT, DSST, PASAT	Mixed	Mixed	✓	✓[a]
	ANAM	Procedural RT	SDMT, DSST, PASAT	✓	✓		✓
	NES2	Symbol Digit	SDMT, DSST, PASAT	✓	✓		
	NES2	Pattern Recognition	SDMT, DSST, PASAT	✓	✓		
	SRT & CRT	SRT	SDMT, DSST, PASAT	✓	✓		✓
		CRT	SDMT, DSST, PASAT	✓	✓		✓
	ANAM	SRT	SDMT, DSST, PASAT	✓	✓		
	Amsterdam Neuropsychological Test	Baseline Speed	Novel	✓	✓		
	Foong Tests	SRT	SDMT, DSST, PASAT		✓		
	Foong Tests	Cued CRT	SDMT, DSST, PASAT		✓		
	Foong Tests	Warned CRT	SDMT, DSST, PASAT		✓		
	Foong Tests	Computerized Symbol Digit Substitution	SDMT, DSST, PASAT		✓		
	Kujala Tests	Number Set	SDMT, DSST, PASAT		✓		
	Kujala Tests	Letter Set	SDMT, DSST, PASAT		✓		
	Kujala Tests	SRT	SDMT, DSST, PASAT		✓		
	Kujala Tests	2-Choice RT	SDMT, DSST, PASAT		✓		
	Kujala Tests	10-Choice RT	SDMT, DSST, PASAT		✓		

(Continued)

Table 3. (Continued)

Functional domain	Battery/test name	Subtest	Comparator(s)	Test–retest reliability	Validity		
					Discriminant/ known groups	Ecological	Concurrent
	Kujala Tests	Statement Verification Test	SDMT, DSST, PASAT		✓		
	Kujala Tests	Subtraction Test	SDMT, DSST, PASAT		✓		
	c-ILT	c-ILT	SDMT, DSST, PASAT		✓		
	GCDT	GCDT	SDMT, DSST, PASAT		✓		
	Denney Tests	Picture Naming Test	SDMT, DSST, PASAT		✓		
	Denney Tests	Rotated Figures Test	SDMT, DSST, PASAT		✓		✓
	Janculjak Tests	Visual Reaction Task	SDMT, DSST, PASAT		✓	n.s.	
	CANTAB	RT	SDMT, DSST, PASAT		✓	n.s.	
	Denney Tests	Remote Associates Test	SDMT, DSST, PASAT		Mixed		✓
	Auto-SDMT	Auto-SDMT	SDMT		✓	n.s.	✓[a]
	CSI	Number Sequencing	Letter Number Sequencing (WMS)				✓[a]
	CSI	Animal Decoding	DSST				✓[a]
	CSI	Symbol Scanning	Symbol Search (WAIS)				✓[a]
	N-Back	0-Back	SDMT, DSST, PASAT	Poor	Mixed		✓[a]
	CSI	Processing Speed Composite	SDMT, DSST, PASAT	✓			✓[a]
	ANAM	Coding Substitution Learning	SDMT or DSST	✓	n.s.		✓[a]
	fMRI-SDMT	fMRI-SDMT	SDMT		n.s.		✓[a]
	CNSVS	Reaction Time	SDMT, DSST, PASAT	Mixed			✓
Working memory	CANTAB	Spatial Working Memory	WMS Symbol Span	✓	✓	✓	
	PVSAT	PVSAT-4 seconds	PASAT	Poor	✓		✓[a]
	PVSAT	PVSAT-3 seconds	PASAT		Mixed	n.s.	✓[a]
	PVSAT	PVSAT-2 seconds	PASAT	✓	✓	n.s.	✓[a]
	NeuroTrax	Working Memory	PASAT		✓		
	Automated Ospan	AOspan	PASAT, Digit Span, Ospan (non-automated)	✓	✓		✓[a]

(Continued)

Table 3. (Continued)

Functional domain	Battery/test name	Subtest	Comparator(s)	Test–retest reliability	Validity		
					Discriminant/ known groups	Ecological	Concurrent
	Sternberg Memory Scanning Task	Sternberg Memory Scanning Task	PASAT, Digit Span, Reading Span		✓	Mixed	n.s.
	VT-SAT	1-Back	PASAT		✓		
	VT-SAT	2-Back	PASAT		✓		✓[a]
	VT-SAT	2-Back at 1-Back Threshold	PASAT		✓		
	Computerized Digit Span	Computerized Digit Span	Digit Span		✓		
	CBB	1-Back	PASAT, WMS Spatial Span, Digit Span	Mixed	Mixed	✓	✓[a]
	N-Back	Overall	PASAT, Digit Span, Reading Span		Mixed		
	N-Back	1-Back	PASAT, Digit Span, Reading Span	Poor	Mixed		✓[a]
	N-Back	2-Back	PASAT, Digit Span, Reading Span	Poor	Mixed		✓[a]
	N-Back	3-Back	PASAT, Digit Span, Reading Span	Mixed	Mixed		Mixed
	CANTAB	Spatial Span	WMS Spatial Span	✓	Mixed		
	Computerized Salthouse Operational Working Memory Test/Keeping Track Task (KTT)	KTT	PASAT or Digit Span		✓	n.s.	✓[a]
	MAT	Verbal Working Memory	PASAT		✓		
	MAT	Episodic Working Memory	WMS Logical Memory Immediate		✓		
	MAT	Figural Working Memory	WMS Spatial Span		n.s.		
	CDR	Quality of Working Memory	PASAT	Mixed	n.s.	✓	Mixed
Episodic memory	NeuroTrax	Memory	CVLT2, BVMTR	✓	✓	✓	✓[a]
	CBB	Continuous PAL Task	BVMTR, 10/36	✓	✓	✓	✓
	CNSVS	Composite Memory	CVLT2, BVMTR	Mixed	✓	✓	✓[a]

(Continued)

Table 3. (Continued)

Functional domain	Battery/test name	Subtest	Comparator(s)	Test–retest reliability	Validity		
					Discriminant/ known groups	Ecological	Concurrent
	CBB	One Card Learning	BVMTR	Poor	✓	✓	✓[a]
	ANAM	Coding Substitution Delayed Recall	WAIS-III DSST Recall	✓	✓		✓[a]
	CANTAB	PALT	BVMTR, 10/36	✓	✓	n.s.	
	Cognivue	Memory	BVMTR, 10/36		✓		✓
	CANTAB	Delayed Matching To Sample	BVMTR learning	Poor	✓	✓	
	CANTAB	Spatial Recognition Memory	10/36, WMS Designs	Mixed	✓		
	Touch Panel Tests	Flipping Cards Game	10/36, WMS Designs		✓	✓	
	CPALT	Recall: Unrelated Items	WMS Verbal Paired Associates, CVLT2		✓		
	CPALT	Delayed Recall	WMS Verbal Paired Associates, CVLT2		✓		
	CSI	Memory Composite	10/36, WMS Designs	✓			✓
	CSI	Memory Cabinet 1	10/36, WMS Designs				✓[a]
	CSI	Memory Cabinet 2	10/36, WMS Designs				✓[a]
	MAT	Episodic Short-Term Memory	WMS Logical Memory I		✓	✓	
	MAT	Figural Short-Term Memory	BVMTR Learning		n.s.		
	MAT	Verbal Short-Term Memory	CVLT2 Learning, RAVLT		n.s.		
	CDR	Quality of Episodic Memory	CVLT2, RAVLT, BVMTR	Mixed	n.s.	✓	✓
	NES2	Paired Associate Learning	WMS Verbal Paired Associates		n.s.		
	CPALT	Recall: Related Items	WMS Verbal Paired Associates, CVLT2		n.s.		

(Continued)

Table 3. (Continued)

Functional domain	Battery/test name	Subtest	Comparator(s)	Test–retest reliability	Validity		
					Discriminant/ known groups	Ecological	Concurrent
Attention	CDR	Power of Attention	Digit Span, BTA, TMT, CPT	✓	✓	✓	✓
	CDR	Continuity of Attention	Digit Span, BTA, TMT, CPT	✓	✓	✓	✓
	NeuroTrax	Attention	Digit Span, BTA, TMT, CPT	✓	✓	✓	✓[a]
	Amsterdam Neuropsychological Test	Focused Attention	Digit Span, BTA, TMT, CPT	✓	✓		
	Amsterdam Neuropsychological Test	Memory Search	Digit Span, BTA, TMT, CPT	✓	✓		
	CBB	Identification	Digit Span, BTA, TMT, CPT	Mixed	✓		✓[a]
	Conners' CPT	CPT	Digit Span, BTA, TMT, CPT	✓	✓		
	NES2	Continuous Performance	Digit Span, BTA, TMT, CPT	✓	✓		
	Janculjak Tests	Visual Attention Task	Digit Span, BTA, TMT, CPT		✓		
	Janculjak Tests	Visual Reaction Task	Digit Span, BTA, TMT, CPT		✓	n.s.	
	MAT	Selective Attention	DKEFS CW, Stroop		✓		
	CANTAB	Rapid visual Processing	CPT		✓	n.s.	
	CNSVS	Complex Attention	DKEFS CW, Stroop, DSST	✓	Mixed		
	Attention Network Test	Overall	Novel	✓	Mixed	✓	n.s.
	Attention Network Test	Alerting	CPT	Mixed	Mixed	n.s.	n.s.
	Attention Network Test	Orienting	Novel	Mixed	Mixed		n.s.
	Attention Network Test	Conflict	Novel	✓	n.s.		n.s.
	TAP	Alertness	CPT		Mixed		
	TAP	Divided Attention	Novel		Mixed		
	TAP	Go-No-Go	DKEFS CW, Stroop		n.s.		
	CSI	Attention Composite	Digit Span, BTA, TMT, CPT	✓			✓[a]
	CNSVS	Simple Attention	Digit Span, BTA, TMT, CPT	Poor			✓
	ANAM	Running Memory CPT Throughput	Digit Span, BTA, TMT, CPT	✓			✓[a]
Executive function	NeuroTrax	Executive Function	DKEFS CW, Stroop	✓	✓	✓	✓

(Continued)

Table 3. (Continued)

Functional domain	Battery/test name	Subtest	Comparator(s)	Test–retest reliability	Validity		
					Discriminant/ known groups	Ecological	Concurrent
	Computerized Stroop Test	Word	DKEFS CW, Stroop		✓	✓	✓
	Computerized Stroop Test	Color	DKEFS CW, Stroop		✓	✓	✓
	Computerized Stroop Test	Color-Word	DKEFS CW, Stroop		✓	✓	✓
	CBB	Groton Maze Learning	NAB Mazes		✓	✓	✓[a]
	NIH Toolbox	Modified Flanker	DKEFS CW, Stroop	✓	✓		✓[a]
	Foong Tests	Computerized Stroop	DKEFS CW, Stroop		✓		
	fMRI Tower of London Test	fMRI Tower of London Test	DKEFS Tower Test		✓		
	Denney Tests	Tower of London	DKEFS Tower Test		Mixed		
	Amieva Computerized Tests	Computerized Stroop	DKEFS CW, Stroop		✓	n.s.	
	CANTAB	Stockings of Cambridge	DKEFS Tower Test		Mixed		
	CNSVS	Cognitive Flexibility	WCST, Stroop, NES Switching Attention	✓	✓		✓[a]
	CNSVS	Executive Function	WCST, Stroop, NES Switching Attention	✓	Mixed		✓[a]
	Amieva Computerized Tests	Go-No-Go	DKEFS CW, Stroop		n.s.	n.s.	
Other/ miscellaneous	NeuroTrax	Visual Spatial	JLO	✓	✓	✓	✓[a]
	CNSVS	Psychomotor Speed	NES Finger Tapping Test	✓	✓	✓	✓
	NeuroTrax	Global Cognitive Score	Novel		✓	✓	✓
	NeuroTrax	Motor Skills	NES Finger Tapping Test	✓	✓		
	ANAM	Index of Cognitive Efficiency	Novel	✓	✓		
	Cognivue	Perception	Novel		✓		✓
	Cognivue	Visuomotor	Novel		✓		✓
	CANTAB	Gambling Task	Iowa Gambling Test		✓		✓
	Visual Search Task	Visual Search Task	Novel		✓	n.s.	

(Continued)

Table 3. (Continued)

Functional domain	Battery/test name	Subtest	Comparator(s)	Test–retest reliability	Validity		
					Discriminant/ known groups	Ecological	Concurrent
	Actual Reality Task	Actual Reality Task	Novel	Mixed	✓		✓
	Kujala Tests	Motor Programming Speed	Novel		✓		
	CANTAB	Information Sampling Task	Iowa Gambling Test		✓		
	CogniFit	Multiple	Multiple				
	Urban DailyCog© and DSDT	Daily functioning and driving	Novel			n.s.	n.s.
	CNSVS	Global Score	Novel				
	Touch Panel Tests	Arranging Pictures Game	Novel		✓	✓	
	Touch Panel Tests	Finding Mistakes Game	Novel		n.s.	✓	
	Touch Panel Tests	Beating Devils game	Novel		n.s.		
	NeuroTrax	Verbal Function	COWAT	Poor	n.s.	✓	
	NeuroTrax	Problem Solving	Raven's Progressive Matrices		n.s.		✓

SSRT: Semantic Search Reaction Time, WMS: Wechsler Memory Scales, WAIS: Wechsler Adult Intelligence Scale, CVLT: California Verbal Learning Test, PAL: Paired Associate Learning, WAIS III: Wechsler Adult Intelligence Scale–Third Edition, SRT: Simple Reaction Time, PALT: Paired Associate Learning Test, CPALT: Computerized Paired Associate Learning Test, RAVLT: Rey's Auditory Verbal Learning Test, BTA: Brief Test of Attention, TMT: Trail Making Test, WCST: Wisconsin Card Sorting Test, JLO: Judgment of Line Orientation, DSDT: Driving Simulator Dual Task, COWAT: Controlled Oral Word Association Test, NAB: Neuropsychological Assessment Battery, DSST: Digit Symbol Substitution Test, CRT: Complex Reaction Time, c-ILT: Computerized Incidental Learning Test, AOspan: Automated Operation Span Task, and DKEFS CW: Delis Kaplan Executive Function System Color Word Interference Test
C-SDMT: Computerized Symbol Digit Modalities Test; CSCT: Computerized Speed Cognitive Test; IPS: Information Processing Speed; PASAT: Paced Auditory Serial Addition Task; CDR: Cognitive Drug Research; PST: Processing Speed Test; CNSVS: Central Nervous System-Vital Signs; CTIP: Computerized Test of Information Processing; CBB: CogState Brief Battery; ANAM: Automated Neuropsychological Assessment Metrics; NES: Neurobehavioral Evaluation System; GCDT: Graded Conditional Discrimination Tasks; CANTAB: Cambridge Neuropsychological Test Automated Battery; CSI: Cognitive Stability Index; PVSAT: Paced Visual Serial Addition Test; VT-SAT: Visual Threshold Serial Addition Task; MAT: The Memory and Attention Test; BVMTR: Brief Visuospatial Memory Test-Revised; CPT: Continuous Performance Test; TAP: Test of Attentional Performance; NIH: National Institute of Health; n.s.: non-significant; Mixed: mixed results from one or more studies showing both significance and non-significance or both poor and adequate reliability; Poor: test–retest reliability below $r = 0.60$.
The symbol '✓' indicates significant results for respective psychometric measure ($p < 0.05$). Blank boxes represent missing/uninvestigated for the respective psychometric measure.
[a]Concurrent validity was with a comparator test of the same domain/cognitive function.

Cognitive processing speed. Of the 44 processing speed tests reviewed, all four psychometric properties were examined for the Computerized Symbol Digit Modalities Test (C-SDMT),[24] Computerized Speed Cognitive Test (CSCT),[28] and NeuroTrax Information Processing Speed (IPS) test.[12]

For test–retest reliability, the C-SDMT demonstrated the highest test–retest coefficient among processing speed tests in MS patients (ICC = 0.97) over a mean interval of 103 days.[24] By comparison, the test–retest reliability for the Processing Speed Test (PST) was 0.88 in MS patients.[40] Most every test evaluated in an MS sample showed acceptable cognitive processing speed reliability (see Table 3 and also Supplementary Tables A and B).

For discriminant validity (MS vs HCs), the average effect size for processing speed tests was $d = 1.03$ (SD = 0.42). The C-SDMT showed medium ($d = 0.76$)[24] to large ($d = 1.48$)[58] effect sizes comparable to the conventional SDMT.[59] The CSCT had similar effect sizes comparing HCs with primary progressive MS ($d = 1.8$, $p < 0.001$) and relapsing remitting MS ($d = 0.80$, $p < 0.01$).[60] Other tests of cognitive processing speed discriminated equally well, including the Computerized Test of Information Processing (CTIP; mean $d = 0.82$, SD = 0.37),[61–65] the PST ($d = 0.75$, $p < 0.001$),[40] and to a lesser extent the Auto-SDMT ($d = 0.68$, $p < 0.01$)[26] and CNSVS Processing Speed subtest ($d = 0.52$, $p = 0.046$).[66]

For ecological validity, EDSS correlated with C-SDMT ($r = 0.35$),[24] NeuroTrax IPS ($r = 0.20$),[67] CNSVS Processing Speed ($r = 0.31$),[66] CTIP (range: $r = 0.39$–0.52),[62] and CBB Detection ($r = 0.46$).[68] EDSS did not correlate with CANTAB Reaction Time[69] and Auto-SDMT.[26] Like the conventional SDMT,[70,71] impairment on the CSCT significantly predicted unemployment among MS patients.[28] No other CNAD cognitive processing speed test was studied in this manner.

For concurrent validity, computerized tests based on SDMT correlated with the conventional version: CSCT ($r = 0.88$),[28] C-SDMT ($r = 0.78$),[72] Auto-SDMT ($r = 0.81$),[26] and PST ($r = 0.80$ and $r = 0.75$).[40,73] SDMT also correlated with CTIP (range: $r = 0.29$–0.40),[62] CBB Detection ($r = 0.40$),[74] and CSI Processing Speed Composite ($r = 0.58$),[13] while the following processing speed tests each correlated significantly with their non-SDMT comparators: CDR Speed of Memory,[10] CSI Number Sequencing,[75] CNSVS Processing Speed,[21] and NeuroTrax IPS.[12]

Working memory. Of the 22 working memory CNAD tests, we found published results on three of the psychometric properties for CANTAB Spatial Working Memory and Paced Visual Serial Addition Test 2 (PVSAT-2).

For test–retest reliability, the PVSAT-2 (ICC = 0.75)[41] had the best consistency in MS patients over a mean of 71.7 days.

Overall, working memory CNADs had moderate discriminant validity with an average effect size of $d = 0.70$ (SD = 0.30). The PVSAT showed smaller effects (range: $d = 0.27$–0.61)[29,41] depending on the rate of stimulus presentation. Except for CANTAB Spatial Span (mean $d = 1.11$)[76] and Automated Ospan Test ($d = 0.90$),[77] other CNAD working memory tests generally demonstrated lower effect sizes. For example, the effects derived from the CDR Quality of Working Memory ($d = 0.20$)[10] and some of the Memory and Attention (MAT) Working Memory subtests (range: $d = 0.05$–0.32)[78] were not statistically significant.

For ecological validity, EDSS was significantly correlated with CDR Quality of Working Memory ($r = 0.48$),[10] CANTAB Spatial Working Memory ($r = 0.43$),[69] CBB One-Back ($r = 0.44$),[68] and Sternberg Memory Scanning Task.[32] However, the Sternberg task did not correlate significantly with EDSS in two other studies.[79,80] EDSS correlations with PVSAT[29] and the Computerized Salthouse Operational Working Memory Test were also non-significant.[79] Ecological validity was unexplored for the remaining working memory tests (listed in Table 3).

In regards to construct validity, validity coefficients with either Paced Auditory Serial Addition Task (PASAT) or Digit Span were as follows: Computerized Salthouse Operational Working Memory Test ($r = 0.51$, $r = 0.57$),[79] CBB One-Back ($r = 0.41$, $r = 0.50$),[68] PVSAT ($r = 0.74$),[29] and the Two-Back level ($r = 0.59$).[81]

Episodic memory. Of the 21 episodic memory tests, all four psychometric properties were established for NeuroTrax Memory and CNSVS Composite Memory.

CNSVS Verbal and Visual Memory demonstrated mixed results for test–retest reliability in HCs.[82] ANAM Coding Substitution Delayed Recall had very high consistency (ICC = 0.88) over 30 days.[83] The NeuroTrax Memory Composite[84] also showed very good reliability with an r value of 0.84. There was a wide range of reliability coefficients for the CDR Quality of Episodic Memory in MS patients, from $r=0.50$ to $r=0.82$, depending on the test–retest interval.[10]

For discriminant/known groups validity, the mean effect size for episodic memory tests was $d=0.70$ (SD = 0.30), and the largest effect size noted was for the CNSVS Composite Memory ($d=1.34$).[85]

For ecological validity, EDSS correlated modestly but significantly with several CNAD memory indices, including CNSVS Composite Memory ($r=0.29$),[86] CBB Continuous Paired Associate Learning (CPAL) Task ($r=0.32$),[68] Touch Panel Tests Flipping Cards Game ($r=0.45$),[50] CDR Quality of Episodic Memory ($r=0.33$),[10] CANTAB Delayed Matching to Sample ($r=0.40$),[69] and MAT Episodic Short-Term Memory ($r=0.21$).[78] The CANTAB Paired Associates Learning Test showed non-significant correlations with EDSS[69,87] and all other episodic memory tests were uninvestigated for this standard.

Four episodic memory tests showed good or excellent concurrent validity: CBB One Card Learning correlated with the Brief Visuospatial Memory Test—Revised (BVMTR)[88] ($r=0.83$),[89] CNSVS Verbal Memory correlated with Rey Auditory Verbal Learning Test[90] ($r=0.54$, $r=0.52$),[21] NeuroTrax Memory Composite correlated with the Selective Reminding Test[91] (range: $r=0.61$–0.65),[12] and CSI Memory Cabinet[75] correlated with Family Pictures subtest from the Wechsler Memory Scales[92] ($r=0.65$, $r=0.61$).[75]

Attention. Of the 23 attention tests reviewed, all four psychometric properties were explored for CDR Power and Continuity of Attention and NeuroTrax Attention.

The test–retest reliability was examined in most CNAD attention tests, but for the most part only in HCs, seldom in MS. The CDR Power of Attention subtest had the strongest test–retest consistency in an MS sample (range: $r=0.86$–0.94).[10]

For discriminating MS patients and HCs, the average effect size for attention was $d=0.81$ (SD = 0.27) with the CANTAB Rapid Visual Processing showing the largest effect size ($d=1.12$).[69]

Ecological validity data were scant as correlation with EDSS or other validators was tested in only a few CNADs. For CDR, Power of Attention correlated with EDSS at $r=0.62$ and the correlation with Continuity of Attention was $r=0.43$.[10] Other EDSS correlations were as follows: NeuroTrax Attention $r=0.26$[67] and Attention Network Test Overall score $r=0.48$.[93] EDSS did not correlate with CANTAB Rapid Visual Processing,[69] Attention Network Test-Alerting,[94] and the Amieva Go-No-Go Test.[95]

Concurrent validity data were similarly sparse. Broadening the construct to include tests such as PASAT, Digit Span, or Trail Making Test,[96] we find that NeuroTrax Attention,[12] CBB Identification,[68,89] CSI Attention Composite,[13] and ANAM Running Memory CPT (Continuous Performance Test) throughput[97] correlate with each of their comparators.

Executive function. A total of 14 executive function CNAD tests were reviewed, with NeuroTrax Executive Function being the only one to be evaluated on all four psychometric properties.

NeuroTrax Executive Function showed good test–retest reliability over 3 weeks to several months ($r=0.80$).[84] The average discriminant effect size of executive function measures was $d=0.99$ (SD = 0.40) with CNSVS cognitive flexibility showing the largest ($d=1.67$).[85]

Ecological validity was demonstrated by significant EDSS correlations with NeuroTrax Executive Function ($r=0.28$),[67] the Computerized Stroop Test (range: $r=0.32$–0.40),[98] and CBB Groton Maze Learning ($r=0.32$).[68] EDSS did not correlate with the Amieva Computerized Stroop,[95] while the remaining executive function measures did not examine ecological validity. The NeuroTrax Global Cognitive Score and Executive Function subtest were studied in relation to employment status, both of which discriminated working and non-working MS patients ($p<0.05$).[99]

A few computerized executive function measures showed moderate associations with similar conventional tests. For instance, the CBB Groton Maze Learning correlated with a maze completion task ($r=0.56$)[100] and the CNSVS measures showed moderate associations with both the Stroop Test and a test of mental set shifting.[21]

Other domains. Briefly, there were several other tests which do not clearly fall within one of the above cognitive domains, most discriminating MS from HCs, but with limited reporting of reliability or ecological validity (Table 3 and Supplementary Tables A and B).

Discussion

In this review, we cast a high bar for designating CNADs as ready for use in MS. We apply the usual standards of psychometric reliability and validity[59,101] and expect that CNAD authors and vendors will publish research focused on MS samples. We recognize that this may present an economic hardship as the CNAD market is highly competitive and vendors are marketing to the wider neurology community. Nevertheless, we maintain that these psychometric standards are important for optimal quality of care, and that research findings should be publicly available, as in conventional NP validation.[102–104]

We find that tests from the CDR, the CBB, NeuroTrax, and CNSVS show acceptable psychometrics. The CSCT, PST, and C-SDMT may be the best single tests for relatively quick screens in busy clinics unable to provide a full NP assessment or carry out longer screening procedures. These tests fulfill the four psychometric criteria (i.e. test–retest reliability, discriminant/known groups validity, ecological validity, and concurrent validity) and are sensitive to MS-related impairment. While the PST lacks a published normative database, we know by personal communication that these data will soon be submitted for peer review. The CSCT and the C-SDMT were used in large healthy samples for comparison. The tests could be made more clinically relevant if age-, sex-, and/or education-specific norms are published to facilitate interpretation of individual results.

There are also notable differences in administration and outcome measures between these versions of the SDMT. The PST uses a manual response on a touchscreen keyboard, while the CSCT, C-SDMT, and traditional SDMT use oral responses recorded by an examiner. Furthermore, the time limits for each of the tests are not identical, as the PST has a 120-second limit, while the CSCT and traditional SDMT both last 90 seconds. The C-SDMT has no time limit as its primary outcome is completion time. These differences prevent valid comparison of raw scores.

Our committee debated the use of the term computerized neuropsychological testing device (CNAD) as published by the American Academy of Clinical Neuropsychology.[4] Likewise, the term paper-and-pencil testing fails to accurately characterize conventional tests. It makes some sense to draw a distinction between tests that present stimuli on a computer screen as opposed to a person speaking to a patient (as in reading a word list or asking a patient to orally list words conforming with a category) or showing them a visual stimulus (as in presenting the Rey figure or interlocking pentagons with the instruction to copy it). Yet there are many shades of gray. The PASAT stimuli are often presented via audio files played on a computer device. Shall we refer to the PASAT as a computerized test? Clearly, the terms computerized and automated refer to a spectrum of technology, ever growing in our effort to improve test accuracy and ease of access.

Relatedly, very few tests have published data on the importance of technician oversight. Supervision could influence performance on several levels. First, it may improve patient motivation to perform well. Second, a technician could provide guidance or clarify instructions for patients with relatively little computer experience or having difficulty understanding a given task. Third, a technician may assist with any technical issues that may arise during test-taking, including malfunctioning hardware or confusion over interacting with a software interface. Recent findings on the CBB and PST suggest that a technician is not necessary in an MS sample.[40,56] However, future research considering other CNADs, particularly those containing more complex tasks, may yield different results, especially for at-home self-administrations. This latter application would necessitate a basic capacity to manage digital platforms such as iPads and the like, as the home-use technology continues to evolve.

For each battery and test, a distinction should be made between the frequency of use in MS research/clinical trials, the *amount* of available psychometric information, and the *quality* of available psychometric information. As shown in Tables 1 and 2, some CNADs were included in numerous publications with MS patients, but as evident in Supplementary Tables A and B, few of these publications aimed specifically to validate the test. In addition, while some tests may have data for reliability and/or validity, the reported coefficients and effect sizes may be low or non-significant. In selecting a CNAD for routine NP assessment or clinical trials in MS, the degree to which a test meets these psychometric categories must be considered.

That said, we acknowledge that some psychometric standards are more important than others. EDSS is not a crucial ecological validity standard as it is notoriously insensitive to CI and physical and mental MS symptoms are generally weakly correlated. Concurrent validity, a process of construct validation, is not as relevant as predicting quality of life and some CNADs assess novel domains/functions for which there are no existing conventional tests. As is evident in Table 3, several CNAD outcomes lack a good comparator and are simply missed by conventional tests. For example, conventional tests seldom measure reaction time, and when they do it is not to the millisecond per stimulus. Another example is the Information Sampling Task

(IST) from CANTAB that requires visual/spatial processing and also decision-making based on perceived probabilities of gain. While not yet validated against an established measure, it may prove valuable within a very narrow sub-area of executive function. Furthermore, construct validation depends on an established metric for the cognitive domain studied and the issue has not been fully examined even for conventional NP tests in MS.

Yet novel CNADs should still possess adequate test–retest reliability and sensitivity before routine clinical application or inclusion in a clinical trial. An interesting issue for sensitivity is the technological limitation preventing CNAD memory tests from evaluating recall, as opposed to recognition memory, the former notably more sensitive in MS.[105,106] Someday CNADs may employ voice recognition or other methods to record and score recall responses, but in this review all of the memory tests use a recognition format. On the other hand, CNADs offer metrics absent from person-administered testing. Many CNADs are automated, which allows for easy administration and often foregoes the need for trained professionals to give instructions, present stimuli, record responses, and score results. In this way, automation can avoid possible bias or error introduced by a psychometrician. Stimuli can also be more readily changed or randomized in CNADs, providing many alternate test forms that may reduce practice effects with repeat testing. Similarly, some CNADs are less subject to ceiling and floor effects because they have the ability to vary the difficulty or presentation of items based on examinee performance.

Computerized tests might particularly benefit the detection of MS-related CI through their improved sensitivity in measuring reaction time and response speed. Declines in cognitive processing speed are the hallmark CI seen in MS.[107] Computerized tests may facilitate the identification of prodromal deficits by capturing minute differences in response time not identified by traditional tests. In contrast to a few raw score indices derived from a conventional SDMT protocol, most CNADs also generate measures of change in accuracy over the course of a task, change in reaction time, and measures of vigilance decrements. In sum, we opine that CNADs are quite good at measuring cognitive processing speed in MS, and their sensitivity and validity in other domains merit further investigation.

Finally, the practicality of CNADs and their cost-effectiveness are frequently cited as reasons to utilize this approach. All cognitive performance tests, be they person- or computer-administered, require certain physical or sensory capacities of patients (e.g. adequate manual dexterity and visual acuity). Conventional, person-administered tests, require skilled trained examiners, are time-consuming, and may be expensive in some cases. Setting aside the potential advantages of a technician (ensuring motivation to perform well, understanding of instructions), we would like to point out that the SDMT costs roughly US$2 per administration and in its oral-response format requires 5 minutes or less. While a human examiner is needed, there is no cost for computer devices, high-speed internet, or paying a vendor for the service (in our experience vendors typically charge US$20 per test). It is interesting that, in a recent investigation of the conventional and computer forms of SDMT,[26] patients reported a preference for the self-administered computer version. If replicated, this perspective may add value to CNADs over conventional tests, much like the PASAT was largely abandoned due to its stressing patients.[59,101] The newly published CMS (Centers for Medicare & Medicaid Services) guidelines for reimbursement add another layer to the discussion. The new CPT code for automated computerized testing reimburses less than US$5 in the United States. Reimbursement rises when a professional becomes involved in the testing process. Thus, the relative value of completely self-administered CNADs is not yet fully recognized by payers, at least in the United States.

Conclusion

Several computerized tests of cognition are available and applied in MS research. As they currently stand, most CNAD batteries and individual tests do not yet demonstrate adequate reliability and validity to supplant well-established conventional NP procedures such as MS Cognitive Endpoints battery (MS-COG), BICAMS (Brief International Cognitive Assessment for MS), or MACFIMS (Minimal Assessment of Cognitive Function in MS). However, some tests (e.g. certain subtests of the CDR, CBB, NeuroTrax, CNSVS, C-SDMT, PST, and CSCT) possess psychometric qualities that approach or maybe even exceed conventional, person-administered tests and can serve as useful screening tools or supplements to full assessments. Further investigations of these CNADs, especially as they relate to ecological measures and patient-relevant outcomes, are needed before widespread implementation with an MS population.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Supplemental Material

Supplemental material for this article is available online.

References

1. Rao SM, Leo GJ, Bernardin L, et al. Cognitive dysfunction in multiple sclerosis. I. Frequency, patterns, and prediction. *Neurology* 1991; 41(5): 685–691.
2. Rao SM, Grafman J, DiGiulio D, et al. Memory dysfunction in multiple sclerosis: Its relation to working memory, semantic encoding, and implicit learning. *Neuropsychology* 1993; 7: 364–374.
3. Benedict RH, Cookfair D, Gavett R, et al. Validity of the Minimal Assessment of Cognitive Function in Multiple Sclerosis (MACFIMS). *J Int Neuropsychol Soc* 2006; 12: 549–558.
4. Bauer RM, Iverson GL, Cernich AN, et al. Computerized neuropsychological assessment devices: Joint position paper of the American Academy of Clinical Neuropsychology and the National Academy of Neuropsychology. *Arch Clin Neuropsychol* 2012; 27: 362–373.
5. Lapshin H, Oconnor P, Lanctt KL, et al. Computerized cognitive testing for patients with multiple sclerosis. *Mult Scler Relat Disord* 2012; 1: 196–201.
6. Korakas N and Tsolaki M. Cognitive impairment in multiple sclerosis: A review of neuropsychological assessments. *Cogn Behav Neurol* 2016; 29(2): 55–67.
7. Moher D, Liberati A, Tetzlaff J, et al. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA statement. *Ann Int Med* 2009; 151: 264–269.
8. Kurtzke JF. Rating neurologic impairment in multiple sclerosis—An Expanded Disability Status Scale (EDSS). *Neurology* 1983; 33: 1444–1452.
9. Fischer JS, Rudick RA, Cutter GR, et al. The Multiple Sclerosis Functional Composite Measure (MSFC): An integrated approach to MS clinical outcome assessment. *Mult Scler* 1999; 5(4): 244–250.
10. Edgar C, Jongen PJ, Sanders E, et al. Cognitive performance in relapsing remitting multiple sclerosis: A longitudinal study in daily practice using a brief computerized cognitive battery. *BMC Neurol* 2011; 11: 68.
11. Snyder PJ, Jackson CE, Petersen RC, et al. Assessment of cognition in mild cognitive impairment: A comparative study. *Alzheimers Dement* 2011; 7(3): 338–355.
12. Achiron A, Doniger GM, Harel Y, et al. Prolonged response times characterize cognitive performance in multiple sclerosis. *Eur J Neurol* 2007; 14(10): 1102–1108.
13. Erlanger DM, Kaushik T, Broshek D, et al. Development and validation of a web-based screening tool for monitoring cognitive status. *J Head Trauma Rehabil* 2002; 17(5): 458–476.
14. Lowe C and Rabbitt P. Test/re-test reliability of the CANTAB and ISPOCD neuropsychological batteries: Theoretical and practical issues. Cambridge Neuropsychological Test Automated Battery. *Neuropsychologia* 1998; 36(9): 915–923.
15. Letz R, Green RC and Woodard JL. Development of a computer-based battery designed to screen adults for neuropsychological impairment. *Neurotoxicol Teratol* 1996; 18(4): 365–370.
16. Ibarra S. Automated Neuropsychological Assessment Metrics. In: Kreutzer JS, DeLuca J and Caplan B (eds) *Encyclopedia of clinical neuropsychology*. New York: Springer, 2011, pp. 325–327.
17. DeSonneville LM, Boringa JB, Reuling IE, et al. Information processing characteristics in subtypes of multiple sclerosis. *Neuropsychologia* 2002; 40(11): 1751–1765.
18. Smith AD, 3rd, Duffy C and Goodman AD. Novel computer-based testing shows multi-domain cognitive dysfunction in patients with multiple sclerosis. *Mult Scler J Exp Transl Clin* 2018; 4(2). DOI: 10.1177/2055217318767458.
19. Shatil E, Metzer A, Horvitz O, et al. Home-based personalized cognitive training in MS patients: A study of adherence and cognitive performance. *Neurorehabilitation* 2010; 26(2): 143–153.
20. Darby D, Maruff P, Collie A, et al. Mild cognitive impairment can be detected by multiple assessments in a single day. *Neurology* 2002; 59(7): 1042–1046.
21. Gualtieri CT and Johnson LG. Reliability and validity of a computerized neurocognitive test battery, CNS Vital Signs. *Arch Clin Neuropsychol* 2006; 21(7): 623–643.
22. Barker-Collo SL. Quality of life in multiple sclerosis: Does information-processing speed have an independent effect. *Arch Clin Neuropsychol* 2006; 21(2): 167–174.
23. Denney DR, Lynch SG, Parmenter BA, et al. Cognitive impairment in relapsing and primary progressive multiple sclerosis: Mostly a matter

of speed. *J Int Neuropsychol Soc* 2004; 10(7): 948–956.

24. Akbar N, Honarmand K, Kou N, et al. Validity of a computerized version of the Symbol Digit Modalities Test in multiple sclerosis. *J Neurol* 2011; 258(3): 373–379.

25. Akbar N, Banwell B, Sled JG, et al. Brain activation patterns and cognitive processing speed in patients with pediatric-onset multiple sclerosis. *J Clin Exp Neuropsychol* 2016; 38(4): 393–403.

26. Patel VP, Shen L, Rose J, et al. Taking the tester out of the SDMT: A proof of concept fully automated approach to assessing processing speed in people with MS. *Mult Scler J* 2018; 25: 1506–1513.

27. Denney DR, Hughes AJ, Elliott JK, et al. Incidental learning during rapid information processing on the Symbol–Digit Modalities Test. *Arch Clin Neuropsychol* 2015; 30(4): 322–328.

28. Ruet A, Deloire MS, Charre-Morin J, et al. A new computerised cognitive test for the detection of information processing speed impairment in multiple sclerosis. *Mult Scler* 2013; 19(12): 1665–1672.

29. Nagels G, Geentjens L, Kos D, et al. Paced Visual Serial Addition Test in multiple sclerosis. *Clin Neurol Neurosurg* 2005; 107(3): 218–222.

30. Lengenfelder J, Bryant D, Diamond BJ, et al. Processing speed interacts with working memory efficiency in multiple sclerosis. *Arch Clin Neuropsychol* 2006; 21(3): 229–238.

31. Sweet LH, Rao SM, Primeau M, et al. Functional magnetic resonance imaging response to increased verbal working memory demands among patients with multiple sclerosis. *Hum Brain Mapp* 2006; 27(1): 28–36.

32. Janculjak D, Mubrin Z, Brinar V, et al. Changes of attention and memory in a group of patients with multiple sclerosis. *Clin Neurol Neurosurg* 2002; 104(3): 221–227.

33. Sternberg S. High-speed scanning in human memory. *Science* 1966; 153(3736): 652–654.

34. Leavitt VM, Lengenfelder J, Moore NB, et al. The relative contributions of processing speed and cognitive load to working memory accuracy in multiple sclerosis. *J Clin Exp Neuropsychol* 2011; 33(5): 580–586.

35. Fan J, McCandliss BD, Sommer T, et al. Testing the efficiency and independence of attentional networks. *J Cogn Neurosci* 2002; 14(3): 340–347.

36. Lazeron RH, Rombouts SA, Machielsen WC, et al. Visualizing brain activation during planning: The Tower of London Test adapted for functional MR imaging. *AJNR Am J Neuroradiol* 2000; 21(8): 1407–1414.

37. Denney DR, Gallagher KS and Lynch SG. Deficits in processing speed in patients with multiple sclerosis: Evidence from explicit and covert measures. *Arch Clin Neuropsychol* 2011; 26(2): 110–119.

38. Kujala P, Portin R, Revonsuo A, et al. Automatic and controlled information processing in multiple sclerosis. *Brain* 1994; 117(Pt 5): 1115–1126.

39. Macniven JAB, Davis C, Ho MY, et al. Stroop performance in multiple sclerosis: Information processing, selective attention, or executive functioning? *J Int Neuropsychol Soc* 2008; 14: 805–814.

40. Rao SM, Losinski G, Mourany L, et al. Processing Speed Test: Validation of a self-administered, IPad®-based tool for screening cognitive dysfunction in a clinic setting. *Mult Scler J* 2017; 23: 1929–1937.

41. Lapshin H, Lanctot KL, O'Connor P, et al. Assessing the validity of a computer-generated cognitive screening instrument for patients with multiple sclerosis. *Mult Scler* 2013; 19(14): 1905–1912.

42. Foong J, Rozewicz L, Chong WK, et al. A comparison of neuropsychological deficits in primary and secondary progressive multiple sclerosis. *J Neurol* 2000; 247(2): 97–101.

43. Conners CK and Sitarenios G. Conners' Continuous Performance Test (CPT). In: Kreutzer JS, DeLuca J and Caplan B (eds) *Encyclopedia of clinical neuropsychology*. New York: Springer, 2011, pp. 681–683.

44. Conners CK, Pitkanen J and Rzepa SR. Conners 3; Conners 2008. In: Kreutzer JS, DeLuca J and Caplan B (eds) *Encyclopedia of clinical neuropsychology*. 3rd ed. New York: Springer, 2011, pp. 675–678.

45. Sfagos C, Papageorgiou CC, Kosma KK, et al. Working memory deficits in multiple sclerosis: A controlled study with auditory P600 correlates. *J Neurol Neurosurg Psychiatry* 2003; 74(9): 1231–1235.

46. Amieva H, Lafont S, Auriacombe S, et al. Inhibitory breakdown and dementia of the Alzheimer type: A general phenomenon. *J Clin Exp Neuropsychol* 2002; 24(4): 503–516.

47. Zimmermann P and Fimm B. A test battery for attentional performance. In: Leclercq M (ed.) *Applied neuropsychology of attention*. London: Psychology Press, 2004, pp. 124–165.

48. Unsworth N, Heitz RP, Schrock JC, et al. An automated version of the operation span task. *Behav Res Methods* 2005; 37(3): 498–505.
49. Zelazo PD, Anderson JE, Richler J, et al. NIH Toolbox Cognition Battery (CB): Validation of executive function measures in adults. *J Int Neuropsychol Soc* 2014; 20(6): 620–629.
50. Kawahara Y, Ikeda M, Deguchi K, et al. Cognitive and affective assessments of multiple sclerosis (MS) and neuromyelitis optica (NMO) patients utilizing computerized touch panel-type screening tests. *Intern Med* 2014; 53(20): 2281–2290.
51. Lamargue-Hamel D, Deloire M, Saubusse A, et al. Cognitive evaluation by tasks in a virtual reality environment in multiple sclerosis. *J Neurol Sci* 2015; 359(1–2): 94–99.
52. Goverover Y, O'Brien AR, Moore NB, et al. Actual reality: A new approach to functional assessment in persons with multiple sclerosis. *Arch Phys Med Rehabil* 2010; 91: 252–260.
53. Adler G, Bektas M, Feger M, et al. Computer-based assessment of memory and attention: Evaluation of the Memory and Attention Test (MAT). *Psychiatr Prax* 2012; 39(2): 79–83.
54. Utz KS, Hankeln TMA, Jung L, et al. Visual search as a tool for a quick and reliable assessment of cognitive functions in patients with multiple sclerosis *PLoS ONE* 2013; 8: e81531.
55. Stricker NH, Lundt ES, Edwards KK, et al. Comparison of PC and iPad administrations of the CogState Brief Battery in the Mayo Clinic Study of Aging: Assessing cross-modality equivalence of computerized neuropsychological tests. *Clin Neuropsychol* 2018; 33: 1102–1126.
56. Cromer JA, Harel BT, Yu K, et al. Comparison of cognitive performance on the CogState Brief Battery when taken in-clinic, in-group, and unsupervised. *Clin Neuropsychol* 2015; 29(4): 542–558.
57. Wojcik CM, Rao SM, Schembri AJ, et al. Necessity of technicians for computerized neuropsychological assessment devices in multiple sclerosis. *Mult Scler* 2018, https://www.ncbi.nlm.nih.gov/pubmed/30465463
58. Hughes AJ, Denney DR, Owens EM, et al. Procedural variations in the Stroop and the Symbol Digit Modalities Test: Impact on patients with multiple sclerosis. *Arch Clin Neuropsychol* 2013; 28(5): 452–462.
59. Benedict RH, DeLuca J, Phillips G, et al. Validity of the Symbol Digit Modalities Test as a cognition performance outcome measure for multiple sclerosis. *Mult Scler* 2017; 23(5): 721–733.
60. Ruet A, Deloire M, Charre-Morin J, et al. Cognitive impairment differs between primary progressive and relapsing-remitting MS. *Neurology* 2013; 80(16): 1501–1508.
61. Walker LA, Cheng A, Berard J, et al. Tests of information processing speed: What do people with multiple sclerosis think about them. *Int J MS Care* 2012; 14(2): 92–99.
62. Hughes AJ, Denney DR and Lynch SG. Reaction time and rapid serial processing measures of information processing speed in multiple sclerosis: Complexity, compounding, and augmentation. *J Int Neuropsychol Soc* 2011; 17(6): 1113–1121.
63. Tombaugh TN, Berrigan LI, Walker LAS, et al. The Computerized Test of Information Processing (CTIP) offers an alternative to the PASAT for assessing cognitive processing speed in individuals with multiple sclerosis. *Cogn Behav Neurol* 2010; 23(3): 192–198.
64. Smith AM, Walker LAS, Freedman MS, et al. Activation patterns in multiple sclerosis on the Computerized Tests of Information Processing. *J Neurol Sci* 2012; 312(1–2): 131–137.
65. Mazerolle EL, Wojtowicz MA, Omisade A, et al. Intra-individual variability in information processing speed reflects white matter microstructure in multiple sclerosis. *Neuroimage Clin* 2013; 2: 894–902.
66. Papathanasiou A, Messinis L, Georgiou VL, et al. Cognitive impairment in relapsing remitting and secondary progressive multiple sclerosis patients: Efficacy of a computerized cognitive screening battery. *ISRN Neurol* 2014; 151379.
67. Golan D, Doniger GM, Wissemann K, et al. The impact of subjective cognitive fatigue and depression on cognitive function in patients with multiple sclerosis. *Mult Scler* 2018; 24(2): 196–204.
68. De Meijer L, Merlo D, Skibina O, et al. Monitoring cognitive change in multiple sclerosis using a computerized cognitive battery. *Mult Scler J Exp Transl Clin* 2018; 4(4), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6293367/
69. Cabeca HLS, Rocha LC, Sabba AF, et al. The subtleties of cognitive decline in multiple sclerosis: An exploratory study using hierarchical cluster analysis of CANTAB results. *BMC Neurol* 2018; 18(1): 140.
70. Morrow SA, Drake A, Zivadinov R, et al. Predicting loss of employment over three years in multiple sclerosis: Clinically meaningful cognitive decline. *Clin Neuropsychol* 2010; 24(7): 1131–1145.
71. Benedict RHB, Drake AS, Irwin LN, et al. Benchmarks of meaningful impairment on

the MSFC and BICAMS. *Mult Scler* 2016; 22(14):1874–1882.

72. Patel VP, Walker LAS and Feinstein A. Deconstructing the symbol digit modalities test in multiple sclerosis: The role of memory. *Mult Scler Relat Disord* 2017; 17: 184–189. DOI: 10.1016/j.msard.2017.08.006.

73. Rudick RA, Miller D, Bethoux F, et al. The Multiple Sclerosis Performance Test (MSPT): An iPad-based disability assessment tool. *J Visual Exp* 2014; 30: e51318.

74. Charvet LE, Shaw M, Frontario A, et al. Cognitive impairment in pediatric-onset multiple sclerosis is detected by the Brief International Cognitive Assessment for Multiple Sclerosis and computerized cognitive testing. *Mult Scler J* 2017; 24: 512–519.

75. Younes M, Hill J, Quinless J, et al. Internet-based cognitive testing in multiple sclerosis. *Mult Scler* 2007; 13: 1011–1019.

76. Foong J, Rozewicz L, Quaghebeur G, et al. Executive function in multiple sclerosis: The role of frontal lobe pathology. *Brain* 1997; 120(Pt 1): 15–26.

77. Grech LB, Kiropoulos LA, Kirby KM, et al. Coping mediates and moderates the relationship between executive functions and psychological adjustment in multiple sclerosis. *Neuropsychology* 2016; 30(3): 361–376.

78. Adler G and Lembach Y. Memory and selective attention in multiple sclerosis: Cross-sectional computer-based assessment in a large outpatient sample. *Eur Arch Psychiatry Clin Neurosci* 2015; 265(5): 439–443.

79. Archibald CJ, Wei X, Scott JN, et al. Posterior fossa lesion volume and slowed information processing in multiple sclerosis. *Brain* 2004; 127(Pt 7): 1526–1534.

80. Janculjak D, Mubrin Z, Brzovic Z, et al. Changes in short-term memory processes in patients with multiple sclerosis. *Eur J Neurol* 1999; 6: 663–668.

81. Parmenter BA, Shucard JL, Benedict RH, et al. Working memory deficits in multiple sclerosis: Comparison between the n-back task and the Paced Auditory Serial Addition Test. *J Int Neuropsychol Soc* 2006; 12(5): 677–687.

82. Littleton AC, Register-Mihalik JK and Guskiewicz KM. Test-retest reliability of a computerized concussion test: CNS Vital Signs. *Sports Health* 2015; 7(5): 443–447.

83. Vincent AS, Roebuck-Spencer TM, Fuenzalida E, et al. Test-retest reliability and practice effects for the ANAM General Neuropsychological Screening Battery. *Clin Neuropsychol* 2018; 32(3): 479–494.

84. Schweiger A, Doniger G, Dwolatzky T, et al. Reliability of a novel computerized neuropsychological battery for mild cognitive impairment. *Acta Neuropsychologica* 2003; 1: 407–413.

85. Papathanasiou A, Messinis L, Zampakis P, et al. Corpus callosum atrophy as a marker of clinically meaningful cognitive decline in secondary progressive multiple sclerosis. Impact on employment status. *J Clin Neurosci* 2017; 43: 170–175.

86. Papathanasiou A, Messinis L, Georgiou V, et al. Cognitive impairment in multiple sclerosis patients: Validity of a computerized cognitive screening battery. *Eur J Neurol* 2014; 21: 258.

87. Cotter J, Vithanage N, Colville S, et al. Investigating domain-specific cognitive impairment among patients with multiple sclerosis using touchscreen cognitive testing in routine clinical care. *Front Neurol* 2018; 9: 331.

88. Benedict RHB, Schretlen D, Groninger L, et al. Revision of the Brief Visuospatial Memory Test: Studies of normal performance, reliability, and validity. *Psychol Assess* 1996; 8: 145–153.

89. Maruff P, Thomas E, Cysique L, et al. Validity of the CogState Brief Battery: Relationship to standardized tests and sensitivity to cognitive impairment in mild traumatic brain injury, schizophrenia, and AIDS dementia complex. *Arch Clin Neuropsychol* 2009; 24(2): 165–178.

90. Rey A. *L'examen Clinique en Psychologie*. Paris: Press Universitaire de France, 1964.

91. Buschke F and Fuld PA. Evaluating storage, retention, and retrieval in disordered memory and learning. *Neurology* 1974; 24(11): 1019–1025.

92. Wechsler D. *Wechsler Memory Scale—Revised manual*. New York: Psychological Corporation, 1987.

93. Ayache SS, Palm U, Chalah MA, et al. Orienting network dysfunction in progressive multiple sclerosis. *J Neurol Sci* 2015; 351(1–2): 206–207.

94. Urbanek C, Weinges-Evers N, Bellmann-Strobl J, et al. Attention Network Test reveals alerting network dysfunction in multiple sclerosis. *Mult Scler* 2010; 16: 93–99.

95. Deloire MSA, Salort E, Bonnet M, et al. Cognitive impairment as marker of diffuse brain abnormalities in early relapsing remitting multiple sclerosis. *J Neurol Neurosurg Psychiatry* 2005; 76(4): 519–526.

96. Reitan RM. Validity of the trail making test as an indicator of organic brain damage. *Percept Motor Skills* 1958; 8: 271–276.

97. Wilken JA, Kane R, Sullivan CL, et al. The utility of computerized neuropsychological assessment of cognitive dysfunction in patients with relapsing-remitting multiple sclerosis. *Mult Scler* 2003; 9(2): 119–127.

98. Lynch SG, Dickerson KJ and Denney DR. Evaluating processing speed in multiple sclerosis: A comparison of two rapid serial processing measures. *Clin Neuropsychol* 2010; 24(6): 963–976.

99. Gudesblatt M, Zarif M, Bumstead B, et al. Multiple sclerosis, cognitive profile and cognitive testing: Predictability of SDMT and computerized cognitive testing in differentiating employment versus unemployment in patients with multiple sclerosis (P5. 207). *Neurology* 2015; 84, https://n.neurology.org/content/84/14_Supplement/P5.207

100. Pietrzak RH, Olver J, Norman T, et al. A comparison of the CogState Schizophrenia Battery and the Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) Battery in assessing cognitive impairment in chronic schizophrenia. *J Clin Exp Neuropsychol* 2009; 31: 848–859. DOI: 10.1080/13803390802592458.

101. LaRocca NG, Hudson LD, Rudick R, et al. The MSOAC approach to developing performance outcomes to measure and monitor multiple sclerosis disability. *Mult Scler* 2018; 24(11): 1469–1484.

102. Erlanger DM, Kaushik T, Caruso LS, et al. Reliability of a cognitive endpoint for use in a multiple sclerosis pharmaceutical trial. *J Neurol Sci* 2014; 340(1–2): 123–129.

103. Benedict RHB, Amato MP, Boringa J, et al. Brief International Cognitive Assessment for MS (BICAMS): International standards for validation. *BMC Neurol* 2012; 12: 55.

104. Benedict RHB, Cookfair D, Gavett R, et al. Validity of the Minimal Assessment of Cognitive Function in Multiple Sclerosis (MACFIMS). *J Int Neuropsychol Soc* 2006; 12: 549–558.

105. DeLuca J, Gaudino EA, Diamond BJ, et al. Acquisition and storage deficits in multiple sclerosis. *J Clin Exp Neuropsychol* 1998; 20(3): 376–390.

106. Rao SM. On the nature of memory disturbance in multiple sclerosis. *J Clin Exp Neuropsychol* 1989; 11(5): 699–712.

107. DeLuca J, Chelune GJ, Tulsky DS, et al. Is speed of processing or working memory the primary information processing deficit in multiple sclerosis? *J Clin Exp Neuropsychol* 2004; 26: 550–562.